

score greater than or equal to the score of the result being printed and is derived by analysts of the total score distribution.

OM nucleic - nucleic search, using sw model

(without alignments)
8025.885 million cell updates/sec

Sequence: 1 ccgcaatgacgcgcac...cgctagccttgcgcg 559

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Pre-processing:	Minimum Match	0%
Post-processing:	Minimum Match	0%

Listing first 45 summaries

Database :

GenEmb1: *

- 1: gb_ba: *
- 2: gb_hhg: *
- 3: gb_in: *
- 4: gb_om: *
- 5: gb_ov: *
- 6: gb_pat: *
- 7: gb_ph: *
- 8: gb_pl: *
- 9: gb_pr: *
- 10: gb_ro: *
- 11: gb_sts: *
- 12: gb_sy: *
- 13: gb_un: *
- 14: gb_vl: *
- 15: em_ba: *
- 16: em_fun: *
- 17: em_hum: *
- 18: em_in: *
- 19: em_mu: *
- 20: em_om: *
- 21: em_ov: *
- 22: em_ov: *
- 23: em_pat: *
- 24: em_ph: *
- 25: em_pl: *
- 26: em_ro: *
- 27: em_sts: *
- 28: em_un: *
- 29: em_vl: *
- 30: em_hhg_hum: *
- 31: em_hhg_in: *
- 32: em_hhg_other: *
- 33: em_hhg_mus: *
- 34: em_hhg_pln: *
- 35: em_hhg_pod: *
- 36: em_hhg_mam: *
- 37: em_hhg_vrt: *
- 38: em_sy: *
- 39: em_hhg_hum: *
- 40: em_hhg_mus: *
- 41: em_hhg_other: *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	559	100.0	559	6	AR073589	AR073589 Sequence
2	559	100.0	1583	6	AX400036	AX400036 Sequence
3	436.2	78.0	1679	9	AK093402	AK093402 Homo sapi
4	435.4	77.9	1726	9	AK094084	AK094084 Homo sapi
5	431.4	77.2	1432	9	AK095965	AK095965 Homo sapi
6	425.2	76.1	1835	9	BC019324	BC019324 Homo sapi
7	425.2	76.1	1844	9	BC004172	BC004172 Homo sapi
8	122.4	21.9	189149	2	AC046171	AC046171 Homo sapi
9	122.4	21.9	194567	2	AC055866	AC055866 Homo sapi
10	113.8	20.4	189149	2	AC046171	AC046171 Homo sapi
11	103	18.4	782	9	HUMPERA	L24804 Human (p23)
12	103	18.4	1490	9	BC003005	BC003005 Homo sapi
13	103	18.4	3342	9	AK057820	AK057820 Homo sapi
14	101.8	18.2	2928	9	AK098211	AK098211 Homo sapi
15	101.4	18.1	127436	6	BD014559	BD014559 Homo sapi
16	101.2	18.1	483	6	BD012892	BD012892 Protein h
17	101.2	18.1	483	23	BD008327	BD008327 Protein h
18	96.4	17.2	1208	10	BC003708	BC003708 Mus muscu
19	96.4	17.2	1510	10	AB024935	AB024935 Mus muscu
20	94.6	16.9	483	10	AF153479	AF153479 Mus muscu
21	94.4	16.9	84558	2	AC122346	AC122346 Rattus no
22	93.4	16.7	134004	9	AC079768	AC079768 Homo sapi
23	93.4	16.7	135342	9	AC106878	AC106878 Homo sapi
24	93	16.6	570	6	AX341269	AX341269 Sequence
25	90.4	16.2	79302	2	AC023898	AC023898 Mus muscu
26	90.4	16.2	139242	10	AL590996	AL590996 Mouse DNA
27	84.8	15.2	468	5	CHKPROGNEC	L24898 Gallus gall
28	83.4	14.9	126452	2	AC118313	AC118313 Rattus no
29	83.4	14.9	146996	2	AC116195	AC116195 Rattus no
30	80.4	14.4	221184	2	AL173315	AL173315 Mus muscu
31	75.4	13.5	3455	9	HSZNF	X56465 Human znf6
32	75.2	13.5	473	6	AX440283	AX440283 Sequence
33	56.6	10.1	1776	6	AX364853	AX364853 Sequence
34	54.2	9.7	128928	9	AC068487	AC068487 Homo sapi
35	54.2	9.7	179599	2	AC013278	AC013278 Homo sapi
36	53.8	9.6	131753	9	AL358790	AL358790 Human DNA
37	53.8	9.6	113758	9	AC009144	AC009144 Homo sapi
38	46.6	8.3	124557	2	AL390834	AL390834 Human DNA
39	46.6	8.3	160737	2	AC026213	AC026213 Homo sapi
40	46.6	8.3	191064	2	AL161745	AL161745 Homo sapi
41	46.6	8.3	203722	2	AC026195	AC026195 Homo sapi
42	43.4	7.8	137329	2	AC095733	AC095733 Rattus no
43	43.4	7.8	200289	2	AC090681	AC090681 Homo sapi
44	43.4	7.8	202235	2	AC117378	AC117378 Homo sapi
45	42	7.5	7218	6	166454	166454 Sequence 14

TITLE	Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,			
JOURNAL	Suzano,S., Nagahari,K., Maunoh,Y., Nagai,K. and Isogai,T.			
REFERENCE	NEO human cDNA sequencing project			
AUTHORS	Unpublished			
TITLE	2 (bases 1 to 1679)			
JOURNAL	Isogai,T. and Yamamoto,J.			
COMMENT	Direct Submission			
	Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7			
	Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan			
	(E-mail:genom@skhri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)			
	NEO human cDNA sequencing project supported by Ministry of			
	Economy, Trade and Industry of Japan: cDNA full insert sequencing:			
	Research Association for Biotechnology (RAB); cDNA library			
	construction: Helix Research Institute (HRI) (supported by Japan			
	Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,			
	HRI, and Biotechnology Center, National Institute of Technology and			
	Evaluation; clone selection for full insert sequencing: HRI and			
	RAB; annotation: HRI and RAB.			
FEATURES	Location/Qualifiers			
SOURCE	1. 1679			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone="TEST12020012"			
	/tissue_type="testis"			
	/clone_lib="TEST12"			
	/note="Cloning vector: PME18SFL3"			
BASE COUNT	372 a	435 c	422 g	450 t
ORIGIN				
Query Match	78.0%	Score 436.2	DB 9	Length 1679;
Best Local Similarity	99.3%	Pred. No.5.8e-112;		
Matches 438:	Conservative	0;	Mismatches 3;	Indels 0;
			Gaps	0;
QY	119	TCGTGTACGTGCACGAATGCCAGATGAGATGGATGTGACAAATGAGATTGATTCATAGC	178	
Db	899	TCCTCCTCAGTCGACGAAGATGCCAGATGGATGGAGTTGTACAAATGATTCATAGC	958	
QY	179	CAAAATGAATCTCCAGAGACTCCACAGATTAAGCGCTCTTCCGCTCTAATTAATCTGTTTGT	238	
Db	959	CAAAATGAATCTCCAGAGACTCCACAGATTAAGCGCTCTTCCGCTCTAATTAATCTGTTTGT	1018	
QY	239	GAGAAATGGAAGAAAAGGTGGCTGGCCGGGCTTCCAAAGAGGATTAACAGCCAGT	298	
Db	1019	GAGAAATGGAAGAAAAGGTGGCTGGCCGGGCTTCCAAAGAGGATTAACAGCCAGT	1078	
QY	299	GTCGCTGTCTGTGACATTTGATTAACGTGAGAGACTGGGAGGGGATGAAGATGAGACT	358	
Db	1079	GTCGCTGTCTGTGACATTTGATTAACGTGAGAGACTGGGAGGGGATGAAGATGAGACT	1138	
QY	359	GGCTCATGTGGAACATTTATGACAGACTTTTGAAGAAGTCAACACCAAGACCTCCAGC	418	
Db	1139	GGCTCATGTGGAACATTTATGACAGACTTTTGAAGAAGTCAACACCAAGACCTCCAGC	1198	
QY	419	TGCCATGATATTTGGATGATGATTTGACAGAGTCCATATATGCAACAAGTAATTAATCT	478	
Db	1199	TGCCATGATATTTGGATGATGATTTGACAGAGTCCATATATGCAACAAGTAATTAATCT	1258	
QY	479	TTCTGTGACGAAGAAGCTGGGAAGCAGCTGTGGCATTTTCCAGTTGTTCTAGAAAGCTA	538	
Db	1259	TTCTGTGACGAAGAAGCTGGGAAGCAGCTGTGGCATTTTCCAGTTGTTCTAGAAAGCTA	1318	
QY	539	GCGCCTAGGCTTTGTGTCAGC	559	
Db	1319	GCGCCTAGGCTTTGTGTCAGC	1339	
RESULT 4				
LOCUS	AK094084	1726 bp	mRNA	linear
DEFINITION	Homo sapiens cDNA FLJ36765 f1s, clone 3NB691000191.			
ACCESSION	AK094084			
VERSION	AK094084.1	GI:21753072		
KEYWORDS	c11g; capping; f1s (full insert sequence).			

SOURCE	Homo sapiens neuroblastoma cell line:NB99 CDNA to mRNA, clone.lib3NB691 clone:3NB691000191.
ORGANISM	Homo sapiens
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 Ninomiyama,K., Magatsuma,M., Kanda,K., Kondo,H., Yokoi,T., Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamanoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagaharil,K., Masuno,Y., Nagai,K. and Isogai,T. NEBO human cDNA sequencing project
TITLE	Unpublished
JOURNAL	2 (bases 1 to 1726)
REFERENCE	Isogai,T. and Yamamoto,J. Direct Submission Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2 Kazusa-Kamatari, Kisarazu, Chiba 293-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-338-52-3975, Fax:81-338-52-3986) NEO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- and 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
FEATURES	Location/Qualifiers 1..1726 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="3NB691000191" /cell_line="NB69" /cell_type="neuroblastoma" /clone_id="3NB691" /note="cloning vector: PME18SFU3"
BASE COUNT	439 a 379 c 533 g 375 t
ORIGIN	
Query Match	77.9% Score 435.4; DB 9; Length 1726;
Best Local Similarity	99.8%; Pred. No. 9.8e-112;
Matches 436; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
OY 1	CCGGAATGGCAGCGCAGACGCCCGGACCTTGTGTACGACAGGCCCATGTATGTCTCA 60
Db 25	CCGGAATGGCAGCGCAGACGCCCGGACCTTGTGTACGACAGGCCCATGTATGTCTCA 84
OY 61	TGAGATTGTTGGTGAGAGACAGCACCGATGTCCAGTCCTTATTGAGATACCGGCATTG 120
Db 85	TGAGATTGTTGGTGAGAGACAGCACCGATGTCCAGTCCTTATTGAGATACCGGCATTG 144
OY 121	TGTTCACTGCAAGAATGCCGATGGATGAGATTGTACAATGAGATTGAGTTCTATGGCA 180
Db 145	TGTTCACTGCAAGAATGCCGATGGATGAGATTGTACAATGAGATTGAGTTCTATGGCA 204
OY 181	AAGTGAACCTCCAAAGACTCCAGATTAAGCGCTCTCCCGCTCTATTACTTTGTTTGA 240
Db 205	AAGTGAACCTCCAAAGACTCCAGATTAAGCGCTCTCCCGCTCTATTACTTTGTTTGA 264
OY 241	GAAATGGAAGGAAAGGTGGCCCTGGCCCGGCTTACCAGAGAGATATCAAAGCAGTGT 300
Db 265	GAAATGGAAGGAAAGGTGGCCCTGGCCCGGCTTACCAGAGAGATATCAAAGCAGTGT 324
OY 301	GCGTGTGTGGAATTGATTAATGAGAGACTGGGAAGGGATGAAGAGATGGAGCTGG 360
Db 325	GCGTGTGTGGAATTGATTAATGAGAGACTGGGAAGGGATGAAGAGATGGAGCTGG 384
OY 361	CTCATGTGGAACATTATGCAAGACTTTTGAAGAAGTTCAGACCAAGAGACTCCACCTG 420

Db 385 CTCATGTGGAATATGACAGCTTTTGAAGAAGTGCAGCACAAGAGACTCCACCTG 444
 QY 421 CCATGGATGATTTGGAT 437
 Db 445 CCATGGATGATTTGGAT 461

RESULT 5
 AK056965
 LOCUS Homo sapiens cDNA FLJ32403 fls, clone SKMUS2000365, weakly similar
 DEFINITION to Human (p23) mRNA.
 ACCESSION AK056965
 VERSION AK056965.1 GI:16552512
 KEYWORDS oligo capping, fls (full insert sequence).
 SOURCE Homo sapiens skeletal muscle cDNA to mRNA, clone_11b:SKMUS2
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 AUTHORS Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K.,
 Arita,M., Musashino,K., Yuuki,H., Hara,H., Sugiyama,T., Irie,R.,
 Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
 Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
 Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
 Matsushima,M., Murakawa,K., Kanehori,K., Sugiyama,A., Kawakami,B.,
 Suzuki,Y., Sugano,S., Nagahara,K., Masuno,Y., Nagai,K. and
 Isogai,T.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1432)
 AUTHORS Isogai,T., Otsuki,T. and Sugiyama,T.
 JOURNAL Direct Submission
 Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
 Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: RAB and
 HRI.

FEATURES
 source Location/Qualifiers
 I. 1432
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="SKMUS2000365"
 /tissue_type="skeletal muscle"
 /clone_11b="SKMUS2"
 /note="cloning vector: PME18FLJ3"
 BASE COUNT 355 a 323 c 328 g 426 t
 ORIGIN

Query Match 77.2%; Score 431.4; DB 9; Length 1432;
 Best Local Similarity 99.8%; Pred. No. 1.3e-110;
 Matches 432; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 127 GGTGAAGAATGCGCATGAGTGTGTACAAATGAGATTGATGGCAAGTGA 186
 Db 39 GGTGAAGAATGCGCATGAGTGTGTACAAATGAGATTGATGGCAAGTGA 98
 QY 187 ACTCCAAGAGACTCCAGATTAAGCGCTCTCCGCTCTATTACTTTTGTGAGAAAT 246
 Db 99 ACTCCAAGAGACTCCAGATTAAGCGCTCTCCGCTCTATTACTTTTGTGAGAAAT 158
 QY 247 GGAAGAAAAAGTGGCTGGCGCGGCTTACCAAGAGATATCAAGCCAGTGTGCTGT 306
 Db 159 GGAAGAAAAAGTGGCTGGCGCGGCTTACCAAGAGATATCAAGCCAGTGTGCTGT 218
 QY 307 CTGTGACTTTGATTAAGTGGAGAGACTGGGAAGGGGATGAAGATGGAGCTGCTCATG 366

Db 219 CTGTGACTTTGATTAAGTGGAGAGACTGGGAAGGGGATGAAGATGGAGCTGCTCATG 278
 QY 367 TGGAACTATTATGACAGCTTTTGAAGAAGTGCAGCACAAGAGACTCCACCTGCCATGG 426
 Db 279 TGGAACTATTATGACAGCTTTTGAAGAAGTGCAGCACAAGAGACTCCACCTGCCATGG 338
 QY 427 ATGATTTGGATGATGATTTGTGACAGTGTGATGATCAACAGTAACTTCTGTGTA 486
 Db 339 ATGATTTGGATGATGATTTGTGACAGTGTGATGATCAACAGTAACTTCTGTGTA 398
 QY 487 CCCTAAGCTGGGAAGGACAGCTGTGCTATTTCACAGTTTCTTGAAGAAGTACGCGCTAG 546
 Db 399 CCCTAAGCTGGGAAGGACAGCTGTGCTATTTCACAGTTTCTTGAAGAAGTACGCGCTAG 458
 QY 547 GCCTTTGTCAGCG 559
 Db 459 GCCTTTGTCAGCG 471

RESULT 6
 BC019324 1835 bp mRNA linear PRI 19-DEC-2001
 LOCUS Homo sapiens, hypothetical protein MGC2744, clone MGC:4371
 DEFINITION IMAGE:2823004, mRNA, complete cds.
 ACCESSION BC019324
 VERSION BC019324.1 GI:17939655
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1835)
 AUTHORS Strausberg,R.
 JOURNAL Direct Submission
 Submitted (13-DEC-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov
 COMMENT Contact: MGC help desk
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Lung, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stolt,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
 Series: IRAL Plate: 9 Row: p Column: 19
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 13376885.

FEATURES
 source Location/Qualifiers
 I. 1835
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /db_xref="LocusID:80755"
 /clone="MGC:4371 IMAGE:2823004"
 /tissue_type="lung, small cell carcinoma"
 /clone_11b="NIH-MGC-7"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 146. 1723

CDS

		/codon_start=1 /product="hypothetical protein MGC2744" /protein_id="AAH19324.1" /db_xref="GI:17939636" /translation="MEFCVEDSTDVHVLIEDHRIYFSCKNADGVELYNEIEFYAKNS KDSODKRSRSTICFVRKWKKEVAMPRLTKEDIKPWLVSVDNDRMDEGDEMLAH VEHIAELKRVSTRPPAMDLDFTTIVVSCPAELOTESNGKREYLSGFQVLED TVLPEGGGQDDRGITINDISLVTRRGEQADHTQPLDPSQVILVYRWMERFEDH MQHSGGHLITAAVDHLEFKLTSWELFRSAIELDTPSMTAEQVAIEQSVNEKIR DRLPVNRVELSLDPEVEOVSGRGLPDHAGPIRVNVEGVDSNMCCGTHTNSLDLO VILKITEGKKRNTNLFELGSRVILKMMERSHGTREKALTALCKGADHVEAVKILLO NSTKILQNNLNLRLDLVHAHSIRNSPDMGVITILHRKEDSPFNITIANEISSE TLFLTVGDEKGGGLFLAGPASYETILGPVAVLELGGKGRFQKATKMSRRM EAQHLDDYISTOSAKE"			
BASE COUNT	491 a	396 c	555 g	393 t	
ORIGIN					
Query Match	76.1%;	Score 425.2;	DB 9;	Length 1835;	
Best Local Similarity	99.3%;	Pred. No. 7.4e-109;			
Matches 427;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;	
QY	8	GGCAGGCGACGACGCCCGGACCTTGCTGTACGACAGCCCATGTATGTTCATGAGATT	67		
DB	94	GGGAAGGACAGACGCCCGGACCTTGCTGTACGACAGCCCATGTATGTTCATGAGATT	153		
QY	68	TTGTGTTGAGACAGACCGCATGTCACGTCCTATTGAGATCACCGCATTTGTTGAG	127		
DB	154	TTGTGTTGAGGACACACCGCATGTCACGTCCTATTGAGATCACCGCATTTGTTGAG	213		
QY	128	CTGCAAGAAATGCCGATGAGATGAGTGTACAAATGAGATGATGATGCCAAAGTGA	187		
DB	214	CTGCAAGAAATGCCGATGAGATGAGTGTACAAATGAGATGATGATGCCAAAGTGA	273		
QY	188	CTCCAAGAGACTCCGAGATTAAGCGCTCTCCGCTCTATTAATCTGTTTGGAGAAATG	247		
DB	274	CTCCAAGAGACTCCGAGATTAAGCGCTCTCCGCTCTATTAATCTGTTTGGAGAAATG	333		
QY	248	GAAGGAAAAGGTGGCGCTGCGCGGCTTACCAAGAGGATATCAAGCAGTGTGCTGTC	307		
DB	334	GAAGGAAAAGGTGGCGCTGCGCGGCTTACCAAGAGGATATCAAGCAGTGTGCTGTC	393		
QY	308	TGTGACTTTGATTAACCTGAGAGACTGGAGAGGGATGAAGAGATGAGCGTCTCATGT	367		
DB	394	TGTGACTTTGATTAACCTGAGAGACTGGAGAGGGATGAAGAGATGAGCGTCTCATGT	453		
QY	368	GGAACATTATGACAGAGCTTTTGAAGAAGTCAAGACCAAGAGACCTCCACCTGCATGGA	427		
DB	454	GGAACATTATGACAGAGCTTTTGAAGAAGTCAAGACCAAGAGACCTCCACCTGCATGGA	513		
QY	428	TGATTTGGAT 437			
DB	514	TGATTTGGAT 523			
RESULT 7					
LOCUS	BC004172	1844 bp	mRNA	linear	PRI 12-JUL-2001
DEFINITION	Homo sapiens, clone MGC:2744 IMAGE:2823004, mRNA, complete cds.				
ACCESSION	BC004172				
VERSION	BC004172.1	GI:13278806			
KEYWORDS	MGC.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 1844)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-MAR-2001) National Institutes of Health, Mammalian				
	Gene Collection (MGC), Cancer Genomics Office, National Cancer				
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,				
USA					
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov				

COMMENT	Contact: MGC help desk Email: cgaps-remail.nih.gov Tissue Procurement: DCTD/DTF cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbio.org contact: amandansystemsbiology.org Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Ketteman and Anuradha Madan			
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.nih.gov Series: IRAL Plate: 2 Row: 1 Column: 7. Location/Qualifiers			
source	1. 1844 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="MGC:2744 IMAGE:2823004" /tissue_type="Lung, small cell carcinoma" /clone_id="NIH_MGC-7" /lab_host="DH10B-R" /note="Vector: pOT87" 155..1732 /codon_start=1 /product="Unknown (protein for MGC:2744)" /protein_id="AAH04172.1" /db_xref="GI:13278807" /translation="MEFCVEDSTDVHVLIEDHRIYFSCKNADGVELYNEIEFYAKNS KDSODKRSRSTICFVRKWKKEVAMPRLTKEDIKPWLVSVDNDRMDEGDEMLAH VEHIAELKRVSTRPPAMDLDFTTIVVSCPAELOTESNGKREYLSGFQVLED TVLPEGGGQDDRGITINDISLVTRRGEQADHTQPLDPSQVILVYRWMERFEDH MQHSGGHLITAAVDHLEFKLTSWELFRSAIELDTPSMTAEQVAIEQSVNEKIR DRLPVNRVELSLDPEVEOVSGRGLPDHAGPIRVNVEGVDSNMCCGTHTNSLDLO VILKITEGKKRNTNLFELGSRVILKMMERSHGTREKALTALCKGADHVEAVKILLO NSTKILQNNLNLRLDLVHAHSIRNSPDMGVITILHRKEDSPFNITIANEISSE TLFLTVGDEKGGGLFLAGPASYETILGPVAVLELGGKGRFQKATKMSRRM EAQHLDDYISTOSAKE"			
CDS				
BASE COUNT	493 a	396 c	560 g	393 t
ORIGIN				
Query Match	76.1%;	Score 425.2;	DB 9;	Length 1844;
Best Local Similarity	99.3%;	Pred. No. 7.4e-109;		
Matches 427;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	8	GGCAGGCGACGACGCCCGGACCTTGCTGTACGACAGCCCATGTATGTTCATGAGATT	67	
DB	103	GGGAAGGACAGACGCCCGGACCTTGCTGTACGACAGCCCATGTATGTTCATGAGATT	162	
QY	68	TTGTGTTGAGACAGACCGCATGTCACGTCCTATTGAGATCACCGCATTTGTTGAG	127	
DB	163	TTGTGTTGAGACAGACCGCATGTCACGTCCTATTGAGATCACCGCATTTGTTGAG	222	
QY	128	CTGCAAGAAATGCCGATGAGATGAGTGTACAAATGAGATGATGATGCCAAAGTGA	187	
DB	223	CTGCAAGAAATGCCGATGAGATGAGTGTACAAATGAGATGATGATGCCAAAGTGA	282	
QY	188	CTCCAAGAGACTCCGAGATTAAGCGCTCTCCGCTCTATTAATCTGTTTGGAGAAATG	247	
DB	283	CTCCAAGAGACTCCGAGATTAAGCGCTCTCCGCTCTATTAATCTGTTTGGAGAAATG	342	
QY	248	GAAGGAAAAGGTGGCGCTGCGCGGCTTACCAAGAGGATATCAAGCAGTGTGCTGTC	307	
DB	343	GAAGGAAAAGGTGGCGCTGCGCGGCTTACCAAGAGGATATCAAGCAGTGTGCTGTC	402	
QY	308	TGTGACTTTGATTAACCTGAGAGACTGGAGAGGGATGAAGAGATGAGCGTCTCATGT	367	
DB	403	TGTGACTTTGATTAACCTGAGAGACTGGAGAGGGATGAAGAGATGAGCGTCTCATGT	462	
QY	368	GGAACATTATGACAGAGCTTTTGAAGAAGTCAAGACCAAGAGACCTCCACCTGCATGGA	427	
DB	463	GGAACATTATGACAGAGCTTTTGAAGAAGTCAAGACCAAGAGACCTCCACCTGCATGGA	522	

COMMENT

Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTP/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadan@systemsbio.org
Anup Madan, Rachel Dickhoff, Jessica Fahney, Stephanie Ford, Julia
Greene, Mark Ketteman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAL Plate: 2 Row: f Column: 7.
Location/Qualifiers

FEATURES

source

1..1844
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:2744 IMAGE:2823004"
/tissue_type="lung, small cell carcinoma"
/clone_lib="NIH-MGC-7"
/lab_host="DH10B-R"
/note="vector: pOTB7"
155..1732

CDS

/codon_start=1
/product="unknown (protein for MGC:2744)"
/protein_id="AAH04172.1"
/db_xref="GI:13278807"

/translation="MERVEDSTVDVHVLIEDHRIYFSCKNADGVELYIEIYFAVNS
KDSODKRSRSTICFVRKWKKEVAPRLTKEDIKPWLVSVDNDRMDEGDEMLAH
VEHIAELKRVSTRKPPAMDLDFTTIVVSCPAELOTESNGKREYLSGFQVLED
TVLPEGGGQDDRGITINDISLVTRRGEQADHTQPLDPSQVILVYRWMERFEDH
MQHSGGHLITAAVDHLEFKLTSWELFRSAIELDTPSMTAEQVAIEQSVNEKIR
DRLPVNRVELSLDPEVEOVSGRGLPDHAGPIRVNVEGVDSNMCCGTHTNSLDLO
VILKITEGKKRNTNLFELGSRVILKMMERSHGTREKALTALCKGADHVEAVKILLO
NSTKILQNNLNLRLDLVHAHSIRNSPDMGVITILHRKEDSPFNITIANEISSE
TLFLTVGDEKGGGLFLAGPASYETILGPVAVLELGGKGRFQKATKMSRRM
EAQHLDDYISTOSAKE"

BASE COUNT

493 a 398 c 560 g 393 t

ORIGIN

Query Match 76.1%; Score 425.2; DB 9; Length 1844;
Best Local Similarity 99.3%; Pred. No. 7.4e-109;
Matches 427; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	8	GGCAGGCGACGACCCCGGACCTTGCTGTACGACAGCCCATGTATGTTCATGAGAT	67
DB	103	GGGAAGGACAGACCCCGGACCTTGCTGTACGACAGCCCATGTATGTTCATGAGAT	162
QY	68	TTGTGTTGAGACAGACCGATGTCACGTCCTATTGAGATCACCGCATTTGTTGAG	127
DB	163	TTGTGTTGAGACAGACCGATGTCACGTCCTATTGAGATCACCGCATTTGTTGAG	222
QY	128	CTGCAAGAAATCCGATGAGATGAGTGTACATAGATGAGTTGATGCCAAAGTGA	187
DB	223	CTGCAAGAAATCCGATGAGATGAGTGTACATAGATGAGTTGATGCCAAAGTGA	282
QY	188	CTCCAGGACTCCGAGATTAAGCGCTCTCCGCTCTATTAATCTGTTTGGAGAAATG	247
DB	283	CTCCAGGACTCCGAGATTAAGCGCTCTCCGCTCTATTAATCTGTTTGGAGAAATG	342
QY	248	GAAGGAAAAGGTGGCTGCGCGGCTTACCAAGAGGATATCAAGCAGTGTGCTGC	307
DB	343	GAAGGAAAAGGTGGCTGCGCGGCTTACCAAGAGGATATCAAGCAGTGTGCTGC	402
QY	308	TGTGACTTTGATTAACCTGAGAGACTGGAGAGGGATGAAGAGATGAGCGTCTCATGT	367
DB	403	TGTGACTTTGATTAACCTGAGAGACTGGAGAGGGATGAAGAGATGAGCGTCTCATGT	462
QY	368	GGAACATTATGACAGAGCTTTTGAAGAAAGTCAAGACCAAGACCTCCACCTGCATGGA	427
DB	463	GGAACATTATGACAGAGCTTTTGAAGAAAGTCAAGACCAAGACCTCCACCTGCATGGA	522

OY 428 TGATTGAT 437
 ||||||||
 Db 523 TGATTGAT 532

RESULT 8
AC046171/c
LOCUS AC046171 189149 bp DNA linear HTG 03-MAR-2001
DEFINITION Homo sapiens chromosome 17 clone RP11-948G15 map 17, WORKING DRAFT
ACCESSION AC046171
VERSION AC046171.3 GI:13184209
KEYWORDS HTG: HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 17, clone RP11-948G15
JOURNAL Unpublished
REFERENCES 2 (bases 1 to 189149)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F., Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G., Buzuk, J., Menes, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C., O'Connor, T., O'Donnell, P., O'Neill, D., Olivar, T., Oliver, J., Peterson, K., Pierre, N., Pisanic, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rotman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaferris, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wymann, D., Ye, W., Young, G., Zainoun, J., Zimmer, A. and Zody, M.
COMMENT Direct Submission
 Submitted (13-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02144, USA
 On Mar 2, 2001 this sequence version replaced g1:7712190.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE Genome Center
JOURNAL Center: Whitehead Institute/ MIT Center for Genome Research
COMMENT Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu

Project Information
 Center project name: L7946
 Center clone name: 948-G_15

Summary Statistics
 Sequencing vector: M13; M77815; 40% of reads
 Sequencing vector: Plasmid; n/a; 60% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 185448 bases at least Q40
 Consensus quality: 186995 bases at least Q30
 Consensus quality: 187536 bases at least Q20
 Insert size: 190000; agarose-efp
 Insert size: 187849; sum-of-contigs
 Quality coverage: 10.0 in Q20 bases; agarose-efp
 Quality coverage: 10.1 in Q20.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 14 contigs. The true order of the pieces
 * is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1116: contig of 1116 bp in length
 * 1117 1126: gap of 100 bp
 * 1127 13252: contig of 2036 bp in length
 * 13253 13352: gap of 100 bp
 * 13353 16335: contig of 3183 bp in length
 * 16336 16635: gap of 100 bp
 * 16636 20967: contig of 4332 bp in length
 * 20968 21067: gap of 100 bp
 * 21068 24711: contig of 3644 bp in length
 * 24712 24811: gap of 100 bp
 * 24812 52351: contig of 27540 bp in length
 * 52352 52451: gap of 100 bp
 * 52452 62074: contig of 9623 bp in length
 * 62075 62174: gap of 100 bp
 * 62175 70797: contig of 8623 bp in length
 * 70798 70897: gap of 100 bp
 * 70898 92634: contig of 21737 bp in length
 * 92635 92734: gap of 100 bp
 * 92735 106944: contig of 14210 bp in length
 * 106945 107045: gap of 100 bp
 * 107046 129146: contig of 22102 bp in length
 * 129147 129246: gap of 100 bp
 * 129247 157583: contig of 28337 bp in length
 * 157584 157683: gap of 100 bp
 * 157684 183142: contig of 25459 bp in length
 * 183143 183242: gap of 100 bp
 * 183243 189149: contig of 5907 bp in length.

FEATURES
source
 1. 189149
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17"
 /clone="RP11-948G15"
 /clone_id="RPC1-11 Human Male BAC"
 1. 11116
 /note="assembly_fragment
 clone_end:SP6
 vector_side:left"
 11217. 13252
 /note="assembly_fragment"
 13353. 16335
 /note="assembly_fragment"
 16636. 20967
 /note="assembly_fragment"
 21068. 24711
 /note="assembly_fragment"
 24812. 52351
 /note="assembly_fragment"
 52452. 62074
 /note="assembly_fragment"
 62175. 70797
 /note="assembly_fragment"
 70898. 92634
 /note="assembly_fragment"
 92735. 106944
 /note="assembly_fragment"
 107045. 129146
 /note="assembly_fragment"
 129247. 157583
 /note="assembly_fragment"
 157684. 183142
 /note="assembly_fragment"
 183243. 189149
 /note="assembly_fragment
 clone_end:T7
 vector_side:right"

BASE COUNT 50053 a 45663 c 44537 g 47593 t 1303 others

ORIGIN	Query Match	21.9%	Score 122.4;	DB 2;	Length 189149;
	Best Local Similarity	99.2%;	0;	Mismatches 1;	Indels 0;
	Matches 123;	Conservative	0;	Indels	Gaps 0;
QY	436	ATGAGATTCGACAGTCTGATGATGACACAGAAATTAACCTTCGTGAGCAGCAAGCT	495		
Db	20448	AGGATGATTCGACAGTCTGATGATGACACAGAAATTAACCTTCGTGAGCAGCAAGCT	20389		
QY	496	GGGAGGACAGCTGTGGCATTTTCCAGTGTGTTCTAGAAAGCAGGCGCCAGCCTTTGTC	555		
Db	20388	GGGAGGACAGCTGTGGCATTTTCCAGTGTGTTCTAGAAAGCAGGCGCCAGCCTTTGTC	20329		
QY	556	AGCG 559			
Db	20328	AGCG 20325			
RESULT 9	AC055866/c	194567 bp	DNA	linear	HTG 06-AUG-2002
LOCUS	AC055866				
DEFINITION	Homo sapiens chromosome 17 clone RP11-376M2 map 17, ***				SEQUENCING
ACCESSION	AC055866				
VERSION	AC055866.14	GI:22123743			
KEYWORDS	HTG; HTGS_PHASe1; HTGS_FULLTOP; HTGS_ACTIVEFIN.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	1 (bases 1 to 194567)				
JOURNAL	2 (bases 1 to 194567)				
REFERENCE	2 (bases 1 to 194567)				
AUTHORS	Barren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,				
	Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedal, F.,				
	Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G.,				
	Campoliano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,				
	Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,				
	Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,				
	Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,				
	Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,				
	Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,				
	Klein, J., Labocque, K., Lamazares, R., Landers, T., Lehoczy, J.,				
	Levine, R., Lien, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,				
	McCarthy, M., McKean, P., McGurk, A., McKernan, K., McPheters, R.,				
	Meldrum, T., Menus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,				
	Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,				
	O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,				
	Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,				
	Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,				
	Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,				
	Testaye, S., Theodore, J., Tirelli, A., Travers, M., Triggillo, J.,				
	Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,				
	Young, G., Zainoun, J., Zimmer, A. and Zody, M.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-APR-2000) Whitehead Institute/MIT Center for Genome				
REFERENCE	Research, 320 Charles Street, Cambridge, MA 02141, USA				
AUTHORS	3 (bases 1 to 194567)				
	Barren, B., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S.,				
	Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,				
	Camara, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,				
	Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,				
	Faro, S., Ferreira, P., Fitzerald, M., Gage, D., Galagan, J.,				
	Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,				
	Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A.,				
	Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,				
	Liu, G., Maclean, C., MacDonald, P., Major, J., Matthews, C.,				
	McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V.,				
	Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Noddu, C., Norman, C.H.,				
	O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T., Peterson, K.,				

```

TITLE
JOURNAL
COMMENT
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnuppach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stoianovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (06-Aug-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 6, 2002 this sequence version replaced gi:21313813.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

-----
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L9812
Center clone name: 376_M_2
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1. 1855: contig of 1855 bp in length
* 11856 11955: gap of 100 bp
* 11956 15736: contig of 3781 bp in length
* 15737 15836: gap of 100 bp
* 15837 31925: contig of 16089 bp in length
* 31926 32025: gap of 100 bp
* 32026 148492: contig of 116467 bp in length
* 148493 148592: gap of 100 bp
* 148593 157642: contig of 9050 bp in length
* 157643 157742: gap of 100 bp
* 157743 167387: contig of 9645 bp in length
* 167388 167487: gap of 100 bp
* 167488 185539: contig of 18052 bp in length
* 185540 185639: gap of 100 bp
* 185640 194567: contig of 8928 bp in length.
Location/Qualifiers
1. 194567
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-376M2"
/culture_lib="RCCT-11 Human Male BAC"
BASE COUNT 51484 a 45817 c 46626 g 49854 t 786 others
ORIGIN
Query Match 21.9%: Score 122.4; DB 2; Length 194567;
Best Local Similarity 99.2%; Pred. No. 2.9e+23;
Matches 123; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CY 436 ATGATGATTGTGACAGTCTGATGATGCACACAAGTAATTAACCTTCCTGTGACGCAAACT 495
I | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 153633 AGGATGATTGTGACAGTCTGATGATGCACACAAGTAATTAACCTTCCTGTGACGCAAACT 153574
CY 496 GGGGAAGCAGCTGTGGCATTTTCCAGTGTTCTTGAAGAAGCTAAGGCCCTTAGCCCTTTGTC 555
I | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 153573 GGGGAAGCAGCTGTGGCATTTTCCAGTGTTCTTGAAGAAGCTAAGGCCCTTAGCCCTTTGTC 153514
CY 556 AGCG 1111
DB 153513 AGCG 153510

```

RESULT 10
AC046171 189149 bp DNA linear HTG 03-MAR-2001
LOCUS Homo sapiens chromosome 17 clone RP11-948G15 map 17, WORKING DRAFT
DEFINITION
AC046171.3 GI:131844209
SEQUENCE 14 unordered pieces.
AC046171.3
HTG: HTGS_PHASE1; HTGS_DRAFT.
VERSION
AC046171.3
KEYWORDS
Homo sapiens.
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 189149) Nussbaum, C. and Lander, E.
Birren, B., Linton, L., Nussbaum, C. and Lander, E.
Homo sapiens chromosome 17, clone RP11-948G15
Unpublished
2 (bases 1 to 189149)
Birren, B., Linton, L., Nussbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G.,
Campiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collimore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howard, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Kartas, A.,
Klein, J., Laroque, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Liu, G., Locke, K., Macdonald, P., Margulis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R.,
Meldrum, J., Menous, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testafaye, S., Theodore, J., Turrelli, A., Travers, M., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (13-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 2, 2001 this sequence version replaced gi:7712190.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WTR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7946
Center clone name: 948.G.15
----- Summary Statistics
Sequencing vector: M13; M77815; 40% of reads
Sequencing vector: Plasmid; n/a; 60% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 185448 bases at least Q40
Consensus quality: 186995 bases at least Q30
Consensus quality: 187536 bases at least Q20
Insert size: 190000; agarose-fp
Insert size: 187849; sum-of-contigs
Quality coverage: 10.0 in Q20 bases; agarose-fp
Quality coverage: 10.1 in Q20.
* NOTE: This is a working draft sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

FEATURES	SOURCE	1	11116: contig of 11116 bp in length
1. 189149	1. 189149	11117	11216: gap of 100 bp
/organism="Homo sapiens"	/organism="Homo sapiens"	11217	13252: contig of 2036 bp in length
/db_xref="taxon:9606"	/db_xref="taxon:9606"	13253	13352: gap of 100 bp
/chromosome="17"	/chromosome="17"	13353	16335: contig of 3183 bp in length
/map="17"	/map="17"	16336	16635: gap of 100 bp
/clone="RP11-948G15"	/clone="RP11-948G15"	16636	20967: contig of 4332 bp in length
/clone.lib="RPC1-11 Human Male BAC"	/clone.lib="RPC1-11 Human Male BAC"	20968	21067: gap of 100 bp
1. 11116	1. 11116	21068	24711: contig of 3644 bp in length
/note="assembly-fragment"	/note="assembly-fragment"	24712	24811: gap of 100 bp
clone_end:566	clone_end:566	24812	52351: contig of 27540 bp in length
vector_side:left	vector_side:left	52352	52451: gap of 100 bp
11217. 13252	11217. 13252	52452	62074: contig of 9623 bp in length
/note="assembly-fragment"	/note="assembly-fragment"	62075	62174: gap of 100 bp
13353. 16335	13353. 16335	62175	70797: contig of 8623 bp in length
/note="assembly-fragment"	/note="assembly-fragment"	70798	70897: gap of 100 bp
16636. 20967	16636. 20967	70898	92635: contig of 21737 bp in length
/note="assembly-fragment"	/note="assembly-fragment"	92635	92734: gap of 100 bp
21068. 24711	21068. 24711	92735	106944: contig of 14210 bp in length
/note="assembly-fragment"	/note="assembly-fragment"	106945	107044: gap of 100 bp
24812. 52351	24812. 52351	107045	129146: contig of 22102 bp in length
/note="assembly-fragment"	/note="assembly-fragment"	129147	129246: gap of 100 bp
52452. 62074	52452. 62074	129247	157583: contig of 28337 bp in length
/note="assembly-fragment"	/note="assembly-fragment"	157584	157683: gap of 100 bp
62175. 70797	62175. 70797	183143	183142: contig of 25459 bp in length
/note="assembly-fragment"	/note="assembly-fragment"	183143	183242: gap of 100 bp
70898. 92634	70898. 92634	183243	189149: contig of 5907 bp in length.
92734. 106944	92734. 106944		
107045. 129146	107045. 129146		
129147. 129246	129147. 129246		
157583. 157683	157583. 157683		
183142. 183143	183142. 183143		
183242. 183243	183242. 183243		
189149. 189149	189149. 189149		
location/Qualifiers	location/Qualifiers		
1. 189149	1. 189149		
/organism="Homo sapiens"	/organism="Homo sapiens"		
/db_xref="taxon:9606"	/db_xref="taxon:9606"		
/chromosome="17"	/chromosome="17"		
/map="17"	/map="17"		
/clone="RP11-948G15"	/clone="RP11-948G15"		
/clone.lib="RPC1-11 Human Male BAC"	/clone.lib="RPC1-11 Human Male BAC"		
1. 11116	1. 11116		
/note="assembly-fragment"	/note="assembly-fragment"		
clone_end:566	clone_end:566		
vector_side:left	vector_side:left		
11217. 13252	11217. 13252		
/note="assembly-fragment"	/note="assembly-fragment"		
13353. 16335	13353. 16335		
/note="assembly-fragment"	/note="assembly-fragment"		
16636. 20967	16636. 20967		
/note="assembly-fragment"	/note="assembly-fragment"		
21068. 24711	21068. 24711		
/note="assembly-fragment"	/note="assembly-fragment"		
24812. 52351	24812. 52351		
/note="assembly-fragment"	/note="assembly-fragment"		
52452. 62074	52452. 62074		
/note="assembly-fragment"	/note="assembly-fragment"		
62175. 70797	62175. 70797		
/note="assembly-fragment"	/note="assembly-fragment"		
70898. 92634	70898. 92634		
/note="assembly-fragment"	/note="assembly-fragment"		
92734. 106944	92734. 106944		
/note="assembly-fragment"	/note="assembly-fragment"		
107045. 129146	107045. 129146		
/note="assembly-fragment"	/note="assembly-fragment"		
129147. 129246	129147. 129246		
/note="assembly-fragment"	/note="assembly-fragment"		
157583. 157683	157583. 157683		
/note="assembly-fragment"	/note="assembly-fragment"		
183142. 183143	183142. 183143		
/note="assembly-fragment"	/note="assembly-fragment"		
183242. 183243	183242. 183243		
/note="assembly-fragment"	/note="assembly-fragment"		
189149. 189149	189149. 189149		
/note="assembly-fragment"	/note="assembly-fragment"		
clone_end:566	clone_end:566		
vector_side:right	vector_side:right		
50033 a 45663 c 44537 g 47593 t 1303 others	50033 a 45663 c 44537 g 47593 t 1303 others		
BASE COUNT	BASE COUNT		
ORIGIN	ORIGIN		
Query Match	Query Match		
Best Local Similarity 98.3%; Pred No. 7.6e-21;	Best Local Similarity 98.3%; Pred No. 7.6e-21;		
Matches 115; Conservative 2; Indels 0; Gaps 0;	Matches 115; Conservative 2; Indels 0; Gaps 0;		


```

Oy 13 GGCACAGCCCGGACCTTGTGTAGACAGAGCCCATGTATGTCATGAGATTGTCG 72
      |||
Db 45809 GGCACAGCCCGGACCTTGTGTAGACAGAGCCCATGTATGTCATGAGATTGTCG 45868
Oy 73 TTGAGACAGACGACCGATGTCACGTGCTTATTGAGATCAGCATTTGTCAGCT 129
      |||
Db 45869 TTGAGACAGACGACCGATGTCACGTGCTTATTGAGATCAGCATTTGTCAGCT 45925

RESULT 11
LOCUS HUMRA 782 bp mRNA linear PRI 31-MAY-1994
DEFINITION Human (p23) mRNA, complete cds.
ACCESSION L24804 L24805
VERSION L24804.1 GI:438651
KEYWORDS progesterone receptor.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 782)
AUTHORS Johnson,J.L., Belto,T.G., Krcso,C.J. and Toff,D.O.
TITLE Characterization of a novel 23-kilodalton protein of inactive
          progesterone receptor complexes
JOURNAL Mol. Cell. Biol. 14 (3), 1956-1963 (1994)
MEDLINE 94158868
PUBMED 8114727
FEATURES
source location/Qualifiers
1..782
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="testis"
1..782
/gene="p23"
233..715
/gene="p23"
/codon_start=1
/protein_id="AA18537.1"
/db_xref="GI:438652"
/translacion="MOPASAKWYDRDYVFIKFEVDSKDVNVNFKSKLPFCGLGGS
DNFKHLEIDLFHCIDPNDKSKHRTDRSIIICLRGEGSGOSMPRLTKRKARKINMLSYD
FNMKWMDSDSDSDMSNDRFSSEMANNNMGDEVDVLPYDGAHDDSDSDDEKMPDLE
"
BASE COUNT 232 a 160 c 197 g 193 t
ORIGIN
Query Match 18.4%; Score 103; DB 9; Length 782;
Best Local Similarity 55.4%; Pred. No. 4.8e-18;
Matches 220; Conservative 0; Mismatches 174; Indels 3; Gaps 1;
FEATURES
source
Oy 1 CCGCATGATGCAAGGACGACCCCGACCTTGTGTAGACAGAGCCCATGTATGTCG 60
      |||
Db 222 CCGCATGATGCAAGGACGACCCCGACCTTGTGTAGACAGAGCCCATGTATGTCG 281
Oy 61 TGGAGTTTGTGTAGAGACAGCAGCATGTCACGTGCTTATTGAGATCAGCCGATG 120
      |||
Db 282 TTGATTTTGTGTAGAGACAGCAGCATGTCACGTGCTTATTGAGATCAGCCGATG 341
Oy 121 TGTGAGCTGCAAGAAATGCC--GATGAGTGGAGTGTGACAAAGATGAGTGTCTAG 177
      |||
Db 342 CATTCAGTTGTCTCGGAGAGAGTGAATTTTAAGCATTTAATGAATGATCTTTTTC 401
Oy 178 CCMAAGTGAATCCCAAGGACTCCAGATTAAGCGCTCTCCCGCTCTATTACTTGTG 237
      |||
Db 402 ACTGATTTGATCCAAATGATTTCCAGCAATAAAGAGACAGATCAATTTATGTTGT 461
Oy 238 TGAGAAATGGAAGAAAAGGTGGCCCTGGCCGCTTACCAAGAGGATATCAAGCCAG 297
      |||
Db 462 TACGAAAAGGAGAAATCTGGCCAGTCATGGCCCAAGGTTAAACAAGAAAGGCAAGCTTA 521
Oy 298 TGTGCTGTCTGTGACATTTGATTAAGAGAGACTGGAAAGGATGAAGATGAGAC 357
      |||

```

```

Db 522 ATTGGCTAGTGTGACACTTCAATTAATTGGAAGAGCTGGAGATGATGATGAAGACA 581
Oy 358 TGGCTCATGTGGAACATTAATGACAGAGCTTTTGAGAA 394
      |||
Db 582 TGTCTAATTTTGTATGCTGCTTCTGAGATGATGAACAA 618

RESULT 12
LOCUS BC003005 1490 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, inactive progesterone receptor, 23 kb, clone MGC:4004
IMAGE:2821965, mRNA, complete cds.
ACCESSION BC003005
VERSION BC003005.1 GI:12804292
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1490)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-2001) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
          Email: cgaaps-remail.nih.gov
          Tissue Procurement: DCTP/DMR
          cDNA Library Preparation: Rubin Laboratory
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Genome Sequence Centre,
          BC Cancer Agency, Vancouver, BC, Canada
          info@bcgsc.bc.ca
          Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
          Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
          Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
          Sen Lee, Victor Ling, Carrie Mathewson, Candice McEavey, Steven
          Ness, Pawan Pandoh, Anna-Liisa Parvahu, Parvaneh Saeedi, Jacqueline
          Schein, Duane Smalins, Michael Smith, Lorraine Spence, Jeff Stott,
          Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
          George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAL Plate: 9 Row: d Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5729963
Location/Qualifiers
1..1490
/organism="Homo sapiens"
/db_xref="taxon:9606"
/db_xref="LocusID:10728"
/clone="MGC:4004 IMAGE:2821965"
/tissue_type="lung, small cell carcinoma"
/clone_1lb="NIH-MGC.7"
/lab_host="DH10B-R"
/note="Vector: pORF7"
107..589
/codon_start=1
/product="inactive progesterone receptor, 23 kb"
/protein_id="AAH03005.1"
/db_xref="GI:12804293"
/translacion="MOPASAKWYDRDYVFIKFEVDSKDVNVNFKSKLPFCGLGGS
DNFKHLEIDLFHCIDPNDKSKHRTDRSIIICLRGEGSGOSMPRLTKRKARKINMLSYD
FNMKWMDSDSDSDMSNDRFSSEMANNNMGDEVDVLPYDGAHDDSDSDDEKMPDLE
"
BASE COUNT 473 a 251 c 305 g 461 t
ORIGIN
Query Match 18.4%; Score 103; DB 9; Length 1490;
Best Local Similarity 54.1%; Pred. No. 5.1e-18;

```


Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

fragments are separated by dashed lines.

ECORI	HindIII	E
-----	-----	-----

SeqDerMap	Fingerprint	SeqDerMap	Fingerprint	SeqDerMap	Fingerprint
5331	5337	1657	1581	9554	9588
2184	2176	449	<800	5671	5780
8055	8191	512	<800	3097	3153
486	<800	2814	2848	2638	2680
106	<800	1247	1202	2036	2036
6378	6427	4695	4691	200	<800
307	<800	3373	3417	2223	2254
681	<800	5553	5549	11143	11078
2173	2176	2961	3007	421	<800
1051	1036	3288	3417	7285	7316
454	<800	2485	2499	1790	1791
3073	3104	3614	3604	3843	3797
422	<800	581	<800	1715	1672
434	<800	1848	1863	2867	2901
516	<800	2839	2848	8796	8836
1238	1238	218	<800	749	750
9061	9092	1063	1046	505	<800
4016	4014	551	<800	1127	1095
2926	2968	2953	3007	14980	15052
241	<800	1918	1863	12590	12516
1194	1150	386	<800	1069	1095
851	851	674	<800	3262	3266
11813	11562	56	<800	7517	7619
4215	4184	3033	3007	4002	3992
4640	4544	5998	5951	3823	3797
5752	5672	1178	1202	194	<800
1840	1834	480	<800	1095	1095
738	720	1120	1046	2051	2036
1288	1238	5428	5320	4139	4147
1410	1349	539	<800	5942	5780
2934	2968	566	<800	3260	3266
75	<800	2081	2047	7746	7619
1757	1754	1023	1046	6116	6107
8771	8753	1633	1581		
4263	4184	966	963		

FEATURES	Location/Qualifiers
source	1. 127436
Query Match	18.1%; Score 101.4; DB 9; Length 127436;
Best Local Similarity	53.9%; Pred. No. 2.3e-17;
Matches 257; Conservative	0; Mismatches 211; Indels 9; Gaps
QY 1	CCGCAATGCGACGCGAGCCGCGACCTTGTGTACGACAGGCCATGTATGTCTCA 60
DB 29798	CCCCGTTACACATGACGAGCTGTCTGCAAGGTGTACGACATGAAGGACATGTCTCA 29739
QY 61	TGAGCTTTTGTTGTGAGACGACCGATGTCCAGCTGTCTTGTAGAGATCACCCATTG 120
DB 29738	TTGAATTTTGTGTGAAGACGTAAAGATGTAAATTTTAAAAATCCAAACTTA 29679
QY 121	TGTCAGCTGCAAGATGCC--GATGAGTGTGATGTGACATGATGATGATGCTATG 177
DB 29678	CATTAGCTGTCTCGAGAGAGATGATTTTAAGCATTTAATGAATGATCTTTTC 29619
QY 178	CCAAAGTGAATCCAAAGACTCCAGATTAAGCGCTCTCCGCTCTATTTACTGTTTG 237
DB 29618	ACTGTATTGATTCACAAAGCATTCACAAAGCATTAAGAAAGACGACATCAATTTATGTTT 29559
QY 238	TGAGAAATGAGAGAAAAGTGGCGCTGCGCCGCTTAACCAAGAGGATATCAAGCCAG 297
DB 29558	TACGAAAGGAGAAATCTGACAGTATGCGCCAGGTTAACTAAAGAAAGGCAAGCTTA 29499
QY 298	TGTGCTGTCTGTGAGACTTTGATTAAGTGTGAGAGAGAGAGGAGTGAAGATGAGAC 357
DB 29498	ATTGCTTAGTGTTCATTTCAATTAATGGAAGAGCTGGGAAGATATTCAGATGAAGACA 29439

